Class 02\_HW

Judy Yang

The columns are order number, the number of a particular type of item, the item, and the price. Rows are each item in an order, and for all orders

tail ./chipotle.tsv 1834 orders

cat ./chipotle.tsv | wc -l   4623 lines

4.

grep "Steak Burrito" ./chipotle.tsv | wc -l

grep "Chicken Burrito" ./chipotle.tsv | wc -l

Steak has 368 lines, Chicken has 553 lines

5.

grep ".\*Chicken Burrito.\*Black" ./chipotle.tsv | wc -l

282

grep ".\*Chicken Burrito.\*Pinto" ./chipotle.tsv | wc -l

105

6.

find . -name \*.tsv

./data/sms.tsv

find . -name \*.csv

./data/airlines.csv

./data/bank-additional.csv

./data/bikeshare.csv

./data/drinks.csv

./data/hitters.csv

./data/imdb\_1000.csv

./data/titanic.csv

./data/ufo.csv

./data/vehicles\_test.csv

./data/vehicles\_train.csv

./data/yelp.csv

7.

find . -name \*dictionary\*

**Command Line Tasks**

1. Look at the head and the tail of **chipotle.tsv** in the **data** subdirectory of this repo. Think for a minute about how the data is structured. What do you think each column means? What do you think each row means? Tell me! (If you're unsure, look at more of the file contents.)
2. How many orders do there appear to be?
3. How many lines are in this file?
4. Which burrito is more popular, steak or chicken?
5. Do chicken burritos more often have black beans or pinto beans?
6. Make a list of all of the CSV or TSV files in the DAT8 repo (using a single command). Think about how wildcard characters can help you with this task.
7. Count the approximate number of occurrences of the word "dictionary" (regardless of case) across all files in the DAT8 repo.
8. **Optional:** Use the the command line to discover something "interesting" about the Chipotle data. Try using the commands from the "advanced" section!